CD33 and *SIGLECL1* Immunoglobulin Superfamily Involved in Dementia

Rendina et al

CD33 and Ig Superfamily in Dementia

Antonella Rendina, PhD, [1] Denise Drongitis, PhD, [1] Aldo Donizetti, PhD, [1] Laura Fucci, PhD, [1] Graziella Milan, MD, [1] Francesca Tripodi, MD, [1] Francesca Giustezza, MS, [1] Alfredo Postiglione, MD, [1] Sabina Pappatà, MD, [1] Raffaele Ferrari, PhD, [1] Paola Bossù, PhD, [1] Antonella Angiolillo, PhD, [1] Alfonso di Costanzo, MD, [1] Massimiliano Caiazzo, PhD, and [1] Emilia Vitale, PhD[AQ1][1]

[1]. From the Institute of Biochemistry and Cell Biology (IBBC), CNR (AR, EV); Department of Biology, University of Naples Federico II (DD, AD, LF); Institute of Genetics and Biophysics (IGB), CNR (DD); Geriatric Center "Frullone" ASL NA1 Centro (GM, FT, FG); Department of Internal Medicine & Surgery, University of Naples "Federico II" (AP); Institute of Bioimaging and Biostructures, CNR (SP), Naples, Italy; Department of Neurodegenerative Disease, University College London, London, UK (RF); Clinical and Behavioral Neurology, IRCCS Fondazione Santa Lucia, Rome (PB); Centro Ricerca e Formazione in Medicina dell'Invecchiamento (CeRMI), Università degli Studi del Molise, Ospedale Cardarelli, Campobasso (AA, AdC), Italy; Department of Pharmaceutics Utrecht, Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands (MC); and Department of Molecular Medicine and Medical Biotechnology, University of Naples Federico II, Naples, Italy (MC)[AQ2].
Emilia Vitale is responsible for the DEMENTIA-BIOBANK.[AQ4]

Send correspondence to: Emilia Vitale, PhD, Institute of Biochemistry and Cell Biology (IBBC), CNR, Naples, Italy; E-mail: emilia.vitale@cnr.it[AQ3]

26022020

19042020

25052020

ABSTRACT

Sialic acid-binding immunoglobulin-type lectins, which are predominantly expressed in immune cells, represent a family of immunomodulatory receptors with inhibitory and activating signals, in both healthy and disease states. Genetic factors are important in all forms of dementia, especially in early onset dementia. *CD33* was recently recognized as a genetic risk factor for Alzheimer disease (AD). Here, we present a 2-generation family with 4 members, the father and the 3 siblings, characterized by an early form of unusual dementia exhibiting a behavioral variant close to behavioral variant frontotemporal dementia phenotype and severe forms of memory loss suggestive of AD. We analyzed the *CD33* gene in this family and identified 10 single nucleotide polymorphisms (SNPs) in a linkage disequilibrium block associated with the disease. We also identified a tag SNP, rs2455069-A>G, in *CD33* exon 2 that could be involved with dementia risk. Additionally, we excluded the presence of *C9orf72* expansion mutations and other mutations previously associated with sporadic FTD and AD. The tag SNP association was also analyzed in selected sporadic AD patients from the same Southern Italy region. We demonstrate that *CD33* and *SIGLECL1* have a significantly increased level of expression in these patients.

INTRODUCTION

Substantial effort has been dedicated by the research community to advance the understanding of dementia. In general, estimates of the prevalence of dementia in people aged 60 years and over range from 4.6% in Central Europe to 8.7% in North Africa and the Middle East. This is similar to all other regional estimates that also fall in a relatively narrow band between 5.6% and 7.6%. An estimated 46.8 million people worldwide were living with dementia in 2015. This number is expected to almost double every 20 years, reaching 74.7 million by 2030 and 131.5 million by 2050 (https://www.alz.co.uk/research/world-report-2018).[AQ7]

The 2 major forms of dementia, Alzheimer disease (AD) and frontotemporal dementia (FTD), have an estimated prevalence in Europe of 5000/100 000 (95% CI, 4.73–5.39) (1) and 15–22/100 000 (2), respectively. Both disease entities are clinically and genetically heterogeneous. AD is age-related, nonreversible and has become a global public health challenge. FTD, a less prevalent form of dementia, consists of a significant conglomerate of familial progressive neurodegenerative disorders (3). Predominant features of FTD are prominent behavioral alterations known as behavioral variant frontotemporal dementia (bvFTD) and distinct temporal syndromes associated with a sporadic non-fluent/agrammatic variant of primary progressive aphasia (3–5). The 2 diseases have considerable overlapping clinical manifestations and are associated with many candidate genes. We hypothesize that there is probable involvement of a common molecular pathway that is yet to be understood.[AQ8]

Low-grade chronic inflammatory responses are present in AD patients and have been associated with activation of inflammatory pathways. This suggests that inflammation could be a potential early marker of the neurodegenerative cascade (6). In this regard, CD33 transmembrane receptor is of particular interest. The 67-kDa transmembrane protein CD33 is known to trigger immune cell-cell interactions and binding of extracellular sialylated glycans on other cells or pathogens. CD33 contains one IgV-set domain followed by one C2-set extracellular domain. The sialic-acid binding region is localized on the IgV-set domain. Additionally, CD33 has an immunoreceptor tyrosine-based inhibitory motif (ITIM) and one ITIM-like domain that facilitates an inhibitory signal to the cells. It has been shown that *CD33* expression is increased in AD brains (7), and that it is expressed in cells of the myeloid lineage and may be part of a disease pathway where tyrosine phosphorylation may occur and could modulate microglial activation and AD risk (8).

Moreover, other genes related to *CD33*, such as the sialic acid binding Ig-like lectins (*Siglecs*), may also be involved in the pathogenesis of FTD and AD. Siglecs are categorized into 2 subgroups: One that includes all of the CD33-related Siglecs, molecules that bear a very high degree of homology with the CD33 molecule, and the other consisting of *Siglecs 5–10*, including *SIGLECL1*, a Siglec-like gene. *Siglecs* represent a gene family that is rapidly evolving. It is hypothesized that this is the reason that the mouse and human CD33 are completely different. They have different intracellular domains, different expression patterns and bind to different ligands. The mouse *CdD33* and human *CD33* apparently have clear divergent roles in regulating phagocytosis, underscoring the importance of studying human *CD33* in the susceptibility of AD. The interpretation of mouse *Cd33* functions thus became less clear and it is now hypothesized that the gene may act only as an ITAM (7, 9).

The family reported herein came to our attention because of the unusual phenotype with both overlapping forms of dementia, FTD and AD. We evaluated the relationship between the previously identified *CD33* single nucleotide polymorphisms (SNPs) (7, 8) and identified the tag SNP rs2455069-A>G in exon 2 in association with the disease in the family. A tag SNP represents a single SNP located in the genomic region with high linkage disequilibrium (LD), and identifies a group of SNPs named for the associated haplotype. The LD block whose length is ~10 282 bp, maps upstream of the gene in a short portion of chromosome 19. Rs2455069 results in the amino acid change Arg69Gly, which may be involved in protein modification. The haplotype has been previously described as being linked to cognitive impairment (10). Here we investigated the role of rs2455069-A>G SNP by high resolution melting (HRM) methods (Supplementary Data Fig. S1). Given the unusual phenotype of the family, we also considered testing the rs143332484-C>T *TREM2* exon 2 genomic variant in all the individuals. In our analysis, only the father was rapidly screened for *C9orf72* mutation. He also underwent exome analyses by a NeuroChip that had an updated version of the NeuroX genotyping platform used for variants associated with neurological disease, including *APP*, *PSEN1*, *PSEN2*, *GRN*, *TREM2*, *CD33*, *MAPT*, and *APOE* genes. We investigated the 3 common allelic variants defined by 2 SNPs rs429358 and rs7412, also known as *ApoE-ε2*, *ApoE-ε3*, and *ApoE-ε4*. In addition to *CD33*, all of the family

members were analyzed in depth for the major candidate genes FTD and AD using progranulin (*GRN*) and *MAPT* analysis.

MATERIALS AND METHODS

Patient Collection and Brain Imaging

We investigated familial and some sporadic patients from a Southern Italy region diagnosed with dementia. All of these patients were evaluated at clinical sites by a team that included geriatricians, social workers, and neuropsychologists trained to select and assess patients using standard testing procedures. Blood was collected by venipuncture and serum was prepared by standard methods and analyzed for plasma homocysteine, vitamin B_{12}, folates, and thyroid hormones. Subjects with cognitive impairment underwent cerebral CT or MRI scans and an extensive neuropsychological battery of investigations. In all but one of the familial cases, an F-18 fluorodeoxyglucose position emission tomography (FDG-PET) study was also performed at an early stage; in 2 subjects, an amyloid PET study was performed at a later stage. We extracted genomic DNA and RNA from blood lymphocytes. This study was conducted under ethical principles stated in the Declaration of Helsinki, as well as with approved national and international guidelines for human research. The Ethics Committee of University of Naples reviewed and approved this study, and signed informed consent was obtained from all participants.

PCR Analyses

Genomic DNA was extracted from whole blood by QIAamp DNA Blood Mini Kit (QIAGEN—Hilden, Germany), in accordance with the manufacturer's protocol. Nucleic acid samples were screened for integrity and purity by 1% agarose gel electrophoresis and concentrations were determined by spectrophotometric analysis (Eppendorf BioSpectrometer and Eppendorf μCuvette G1.0, Hamburg, Germany). PCR primers for *CD33*, *GRN*, and *MAPT* genes were designed using the Primer3 tool in order to analyze by PCR amplification the genomic regions encompassing the genes of interest (Supplementary Data Tables S1–S3) and the SNPs within the genes. We used the Sanger sequencing method to obtain reference sequences and to confirm our genomic data analyses. PCR amplifications were performed on a SureCycler 8800 Thermal Cycler (Agilent, Santa Clara, CA).

C9orf72 Repeat Expansion

The hexanucleotide repeat in *C9orf72* was screened only in the father, patient H251. Repeat primed PCR was used as previously described (11, 12). The same patient, H251, was also successfully genotyped by means of the NeuroArray (Table 1) (13). *C9orf72* was analyzed as follows: After PCR cycling, fragment length analysis was performed on an ABI 3730 genetic analyzer (Applied Biosystems, Foster City, CA) and the fragments were visualized using the GeneMapper software. Positive and negative controls were always included as reference. Although this method is not able to determine the exact number of repeats, it can detect repeat numbers up to a maximum of $\sim$60–70. However, the method is able to discriminate repeat counts from $1 \leq n \leq \sim 70$. Therefore, it recognizes repeat counts ≥ 30, which are generally accepted as predictive of the presence of long expansions (i.e. in a pathogenic range).

TABLE 1. NeuroArray Genes Analyzed for Patient H251[AQ10]

1. Gene	2. Clinical Phenotype	3. Chromosomal Location	4. Mutation Reported Position (GRCh37) and Amino Acid Change	5. Variant ID	6. H251 Haplotypes	7. H251 Patient Carrier Status
APP	Alzheimer disease	21q21.3	•1. Chr21:27269905 G>A; E682K •2. Chr21: 27264170 C>G; A692G •3. Chr21:27264167 A>G; E693G •4. Chr21:27264096 G>A; V717I	•1. i5039800 •2. rs63750671 •3. rs63751039 •4. rs63750264	•1. G/G •2. C/C •3. A/A •4. G/G	No mutations
PSEN1	Alzheimer disease	14q24.2	•1. Chr14:73640278 T>C; Y115H •2. Chr14:73640350 A>G; M139V •3. Chr14:73640373 G>C; M146I •4. Chr14:73640371 A>G; M146V	•1. rs63749962 •2. rs63751037 •3. rs63750391 •4. rs63750306	•1. T/T •2. A/A •3. G/G •4. A/A	No mutations
PSEN2	Alzheimer disease-Frontotemporal dementia	1q42.13	Chr1: 227071518 C>T; A85VChr1: 227073304 A>T; N141IChr1: 227076655 A>G; Y231CChr1: 227076678 A>G; M239V	rs63750048 rs63750215 rs200754713 rs28936379	C/CA/AA/AA/A	No mutations
APOE	Alzheimer disease	19q13.32	•1. Chr19:45411941 T>C; C112R •2. Chr 19:45412079 C>T; R158C	•1. rs429358 •2. rs7412	•1. T/T •2. C/T	No mutations
TREM2	Alzheimer disease-Frontotemporal dementia	6p21.1	•1. Chr6:41129295 C>T; Q33X •2. Chr6:41129207 G>A; R62H	•1. rs104894002 •2. rs143332484	•1. C/C •2. G/G	No mutations

1. Gene	2. Clinical Phenotype	3. Chromosomal Location	4. Mutation Reported Position (GRCh37) and Amino Acid Change	5. Variant ID	6. H251 Haplotypes	7. H251 Patient Carrier Status
CD33	Cognitive decline	19q13.41	Chr19:51728641 **A>G; R69G**	***tag SNP rs2455069**	**A/G**	**Mutation**
C9ORF72	Amyotrophic lateral sclerosis, frontotemporal dementia	9p21.2	Chr9: 27573522-27573544; $(G4C2)_{EXP}$	rs143561967	G4C2	No mutation
MAPT	Frontotemporal dementia	17q21.31	•1. Chr17:44095989 G>A; G335S •2. Chr17:44101376 G>A G>C; G389R	•1. rs63750095 •2. rs63750512	•1. G/G •2. G/G	No mutations
GRN	Frontotemporal dementia	17q21.31	Chr17:42427083 T>C; C105R	rs63750441	T/T	No mutations

NeuroArray Genes analyzed on patient H251 by Illumina Infinium platform screening. Column 1, gene name; Column 2, clinical phenotype associated with the specified genes in column 1; Column 3, chromosomal location; Column 4, mutation reported in previous studies associated with disease and genome position according to Homo sapiens (human) genome assembly *GRCh37* (hg19), from Genome Reference Consortium [GCA_000001405.1 GCF_000001405.13] (GRCh37—hg19—Genome—Assembly—NCBI www.ncbi.nlm.nih.gov > assembly), and corresponding amino acid change; Column 5, variant SNP ID number or accession number used to refer to human-specific variants; Column 6, patient's H251 SNP allele genotype revealed in our analysis; Column 7, H251 patient carrier status resulting from our analysis. *CD33 tag SNP rs2455069 is the only mutation resulting in our patient.

NeuroArray Genotyping

Patient H251 was also successfully genotyped by means of the NeuroArray on the Illumina Infinium platform (Illumina, San Diego, CA) (Table 1) (13). The array contains 306 670 highly informative tagging SNPs and 179 467 custom disease-associated variants covering neurodegenerative diseases and was used for identification of high-throughput genome-wide variants. Genotyping call rates ≥ 95% and data set markers with GenTrain scores ≥0.7 were used for quality check. Also, the top set of markers (missing genotype rate 95%, Hardy Weinberg's disequilibrium ≥1e-10 midp and minor allele frequency [MAF] ≥0.01) was used to assess ancestry via standard principal component analysis and HapMap Phase 3 release version 3 (ftp://ftp.hgsc.bcm.tmc.edu/HapMap3-ENCODE/HapMap3/HapMap3v3).

RNA Isolation From Leucocyte Samples and Reverse Transcription Reactions

Total RNA was isolated from selected leucocyte samples using TRIzol Reagent (Invitrogen, Milan, Italy). Samples were rapidly homogenized in TRIzol reagent, incubated for 5 minutes at room temperature, chloroform was added, samples were vortexed for 2 minutes and centrifuged for 15 minutes at 4°C (12 000g). Isopropyl alcohol was added and samples were briefly vortexed and incubated for 10 minutes at room temperature, followed by centrifugation for 10 minutes at 4°C (12 000g). Nucleic acid precipitates were washed in 75% ethanol and centrifuged for 5 minutes at 4°C (7500g). The RNA pellets were allowed to air dry and were resuspended in nuclease-free water. Efficient removal of contaminant DNA was performed via DNAse treatment (DNAse RNAse-free, Ambion, Austin, TX) according to the manufacturer's instructions. RNA was then quantified using a NanoDrop 1000 and checked by denaturant agarose gel electrophoresis. One microgram of RNA was reverse transcribed into cDNA using 200 U of SuperScript III reverse transcriptase (Invitrogen, Milan, Italy) in a final volume of 20 μL. RNA was incubated for 5 minutes at 65°C in a total reaction volume of 13 μL with random hexamer primers and dNTPs. Samples were chilled on ice for at least 1 minute. A cDNA synthesis buffer was added to the reaction and the samples were incubated for 2 minutes at 20°C. Reverse transcriptase was added to the reactions and incubated for 10 minutes at 25°C, followed by 42°C for 50 minutes, and 70°C for 15 minutes.

Quantitative Real-Time PCR

In cases when white blood cells were available, quantitative real-time PCR (qPCR) reactions were performed to determine if they carried the heterozygote tag SNP haplotype A/G (Tables 2 and 3). The subjects tested were H274, H299 (patients with severe AD phenotype carrying the tag SNP); H247, H217, H188, H251 (D8 family bvFTD members carrying the tag SNP); and 3 healthy controls carrying the tag SNP and H252, the healthy mother of the family. The reactions were done in technical duplicates using SYBR green Master Mix in the 7500 HT real-time PCR system (Applied Biosystems). The reaction mixture contained 100–300 ng of cDNA template and 4 pmol/μL of each forward and reverse primer in a final volume of 15 μL. For all experiments, the PCR cycle parameters were 10 minutes at 95°C, followed by 40 cycles of 95°C for 15 seconds and 60°C for the detection of total *CD33* transcripts (*CD33* EX_4-5) or 63°C for discrimination between *CD33* transcript variants (*CD33* EX_1-2 and *CD33* EX_1-2M) for 1 minute. A melting curve analysis was conducted to assess the uniformity of product formation, amplification of non-specific products and primer dimer formation. Primers for qPCR were designed using the Primer3 tool and 18S was used as a reference gene. The primer sequences used are specified in Supplementary Data Table S4. Four pairs of primers were used to detect *CD33* transcript levels, specifically *CD33* EX_4-5 amplified total *CD33* transcript, *CD33* EX_1-2 and *CD33* EX_1-2M used to detect *CD33* transcripts with base variants in exon 2 region (C or T) and *CD33* SPL amplified *CD33* isoform without exon 2 and *SIGLECL1*. Gene expression was determined using the $2e^{-\Delta\Delta CT}$ method and mRNA ratio was calculated using the $2e^{-\Delta CT}$ method. Results are expressed as mean ± standard deviation. Statistical analysis on qPCR data was carried out by a 2-tailed *t*-test (Microsoft Excel software). Results were considered to be statistically significant in case of $p < 0.05$.

TABLE 2. Diagnoses of D8-Behavioral Variant Frontotemporal Dementia Family

1. ID	2. Diagnosis	3. Patients	4. Year of Birth
H188	Behavioral and memory problems	Male son	1971
H217	Behavioral and memory problems	Female daughter	1970
H247	Behavioral and memory problems	Male son	1968
H251	Behavioral	Affected father	1942
H252	NA	Unaffected mother	1947

Column 1, ID; Column 2, clinical diagnosis; Column 3, patient status; Column 4, year of birth. NA, not applicable.

TABLE 3. Patients Investigated by Gene Expression Analyses

1. ID	2. Diagnosis	3. Status	4. Year of Birth	5.Haplotype Tag SNP's rs2455069
H188	Behavioral and memory	D8 family male son	1971	A/G
H274	AD	Male	1958	A/G
H299	AD	Female	1944	A/G

Column 1: ID; Column 2: clinical diagnosis; Column 3: status; Column 4: year of birth; Column 5: haplotype tag SNP's rs2455069 characterization.

RESULTS

We identified an unusual family with 4 affected individuals, including the father H251, the daughter H217, who is the proband, and her 2 brothers, H188 and H247 (Fig. 1; Table 2). They all exhibited apathy or inertia with memory impairment at an early stage, a phenotype characterized mostly by features of bvFTD with executive-predominant deficits on neuropsychological profiling, and relative sparing or less involvement of episodic memory characteristic of AD.

FIGURE 1. D8 disease family pedigree showing segregation of *CD33* risk haplotype (RefSeq NM_001772) associated with cognitive decline. The SNP block is shown in linkage disequilibrium (LD). Circles represent females; squares represent males. Filled symbols represent affected individuals; the open symbol indicates the unaffected individual. The arrow points to the proband, the daughter II-2 H217. The SNP rs2455069 (10th SNP in the haplotype of affected individuals) is the tag SNP of the LD block that we identified through bioinformatics tools (https://snpinfo.niehs.nih.gov/-LD Tag SNP selection).
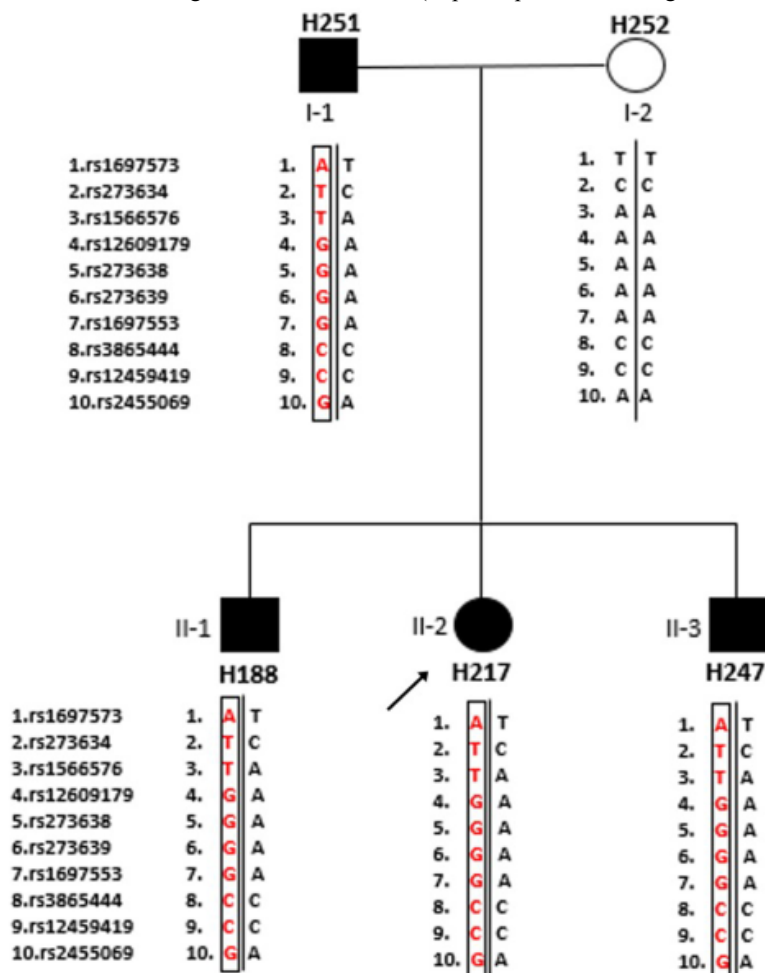
The female patient, H217, age 44, was evaluated because of a 3-year history of behavioral (irritability, agitation, and dysphoria) and cognitive symptoms (progressive memory and executive impairment). Her global standard cognitive examination, Mini-Mental State Examination (MMSE), and the global evaluation of her cognitive functions through Milan Overall Dementia Assessment (MODA) yielded no deficiency scores, even after correcting for age, education, and impairment of functions. The global evaluation of her frontal functions through the Frontal Assessment Battery (FAB) did not reveal any cognitive compromises, but her inhibitory control of automatic replies (Stroop test) (14), and her visual, spatial planning and organizational skills (The Rey-Osterrieth complex figure test), as evaluated by specific frontal tests, were deficient. In addition, the short-term auditory-verbal memory (verbal word span) values were below the standard. Her short-term visual-spatial memory (spatial span) and auditory memory-short term verbal (verbal span figures) recorded values at the limits of normal. The instrumental and basic activities of daily living were minimally compromised. The diagnosis was frontotemporal mild cognitive impairment (15). FDG-PET images showed a moderate reduction of the tracer uptake in the lateral and medial precuneus, more evident on the right side, and in the superior dorsolateral and medial frontal cortex, bilaterally (Supplementary Data Fig. S2). At 1-year follow up the MODA values, anterograde auditory-verbal short- and long-term memory (15 words of Rey) scores and short-term visual-spatial memory (spatial span) and the auditory memory-short term verbal (verbal span figures) values were below the standard values. The cognitive deficits, including the FAB, were further worsened at the follow up of 2019 (Supplementary Data Table S5A, B). Amyloid-PET was negative for amyloid accumulation in the cerebral cortex (Supplementary Data Fig. S2).

The younger brother, H188, was evaluated for the first time at the age of 43 years because his wife noted that he exhibited depression, dysphoria, irritability, and lability. At the clinical interview, the patient was cooperative, adequately understood all the tests administered and was able to perform the entire assessment delineated by the diagnostic protocol. He was well oriented to time and place and had a well-preserved family orientation. His language was fluent and adequately communicative. The global assessment tests MMSE and MODA yielded no pathological values after correction for age and education. The global examination of executive functions through the FAB also did not demonstrate pathological values. At this time, FDG-PET revealed a mildly reduced uptake in the superior dorsolateral and medial frontal cortex bilaterally and in the superior parietal cortex, which was more evident on the left side (Supplementary Data Fig. S2). At follow-up, tests evaluating individual functions exhibited pathological values in short-term spatial memory tests (span spatial). We noted values at the limits of normal in the ability of lexical research by letter (phonological verbal fluency). He exhibited mild problems in performing complex functional tasks that he had performed normally in the past, such as paying bills, preparing meals, or shopping. These deficiencies were characterized as cognitive impairment with behavioral and memory disorders. The clinical and neuropsychological evaluation showed a worsening of cognitive decline and frontal behavioral symptoms with development of dementia. The verbal auditory episodic memory test, executive and problem-solving test (Wisconsin Card Sorting Test) and visuospatial and executive test (Rey-Osterrieth complex figure test copy and recall) were pathological (Supplementary Data Table S5A, B).

The older brother, H247, was initially referred for evaluation at the age of 46 years with reported cognitive function deficits with prevalent memory involvement associated with behavioral disorders (irritability, agitation, and aggressiveness). The neuropsychological evaluation confirmed pathological scores in memory, executive, attention, planning and problem-solving tests (Stroop test, Rey-Osterrieth complex figure test copy). The global cognitive examination (MODA) yielded values below the norm, even after correction for age and education, highlighting his impaired functions. FDG-PET demonstrated a more widespread reduced uptake in the frontal cortex, dorsolateral and bilateral medial regions, as well as in the left parietal, lateral and medial precuneus bilaterally, more evident on the left. At follow-up (Supplementary Data Table S5A, B), he presented with worsening cognitive decline and functional autonomy in basic and instrumental activity of daily living. He exhibited frontal behavioral symptoms of apathy, socially inappropriate behavior, environmental dependence, perseveration and confabulation (Supplementary Data Table S5A, B).

The father, H251, had his first evaluation at age 72 because he exhibited cognitive decline and frontal behavior symptoms (apathy, socially inappropriate behavior, disinhibition, a loss of empathy, and anosognosia). The neuropsychological evaluation indicated pathological scores in memory, attentional-executive, visual and spatial planning and organization capacity, planning and problem-solving tests (spatial span test, Trial Makin Test, Stroop Test, the clock drawing test, Rey-Osterrieth complex figure test copy and recall, Raven matrices). His global cognitive examination

(MMSE and MODA) yielded values below normal even after correction for age and education, and demonstrated global cognitive decline. The instrumental and basic activities of daily living were compromised. The diagnosis was bvFTD and a multidomain cognitive deficit (16, 17) (Table 2). Neuropsychological testing scores are shown in Supplementary Data Table S5A, B. Amyloid-PET did not show increased load in the cerebral cortex (Supplementary Data Fig. S2).

Members of the entire family, including the nonaffected mother, H252, were analyzed for candidate genes including the *APP*, *PSEN1*, *PSEN2*, *GRN*, *TREM2*, *CD33*, *MAPT*, and *APOE* genes. All the family members were analyzed in depth for *CD33* as well as for the major candidate genes FTD and AD by means progranulin (*GRN*) and MAPT. We analyzed all the exons and intron exon bounding sequences for the selected candidate genes by Sanger sequences (Supplementary Data for primers and methods). The father alone was screened for *C9orf72* hexanucleotide repeat using repeat primed PCR. The results revealed a <30 repeat count indicating the absence of mutations. He also underwent NeuroArray genotype analysis. Although a significant number of polymorphic variations were identified, we excluded the presence of other mutations that could have been responsible for the disease (data are available upon request) (11–13). Gene analyses demonstrated an association with *CD33* SNPs block in LD, which may be associated with the unusual disease phenotype in this family (Figs. 1 and 2; Table 2).

FIGURE 2. D8 disease family rs2455069-A>G in electropherograms. Rs2455069 SNP electropherograms obtained by Sanger sequencing method of patients where genomic DNA was amplified by primer set FW: 5′-ACT ACG ACA AGA ACT CCC CAG-3′ and RV: 5′-CCC TGA GTC TCC TCC TGT ACT-3′. The arrows on each electropherogram identify the genotype for the tag Rs2455069 SNP. The sequences confirm that the affected individuals, H188, H217, H247, and H251, have heterozygote genotype A/G while the unaffected mother, H252, has homozygote genotype A/A.
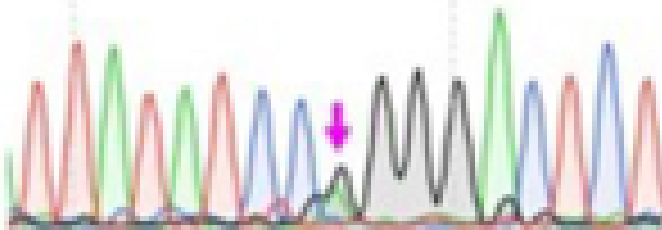
H 188

T T A T A T C C **R** G G G A C T C T
[A/G]
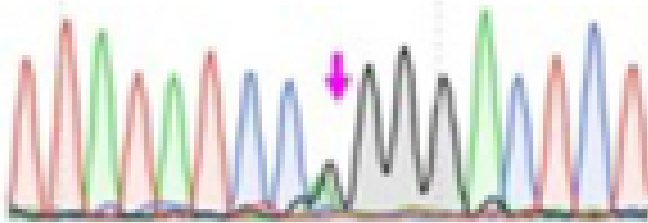
H 217

T T A T A T C C **R** G G G A C T C T
[A/G]
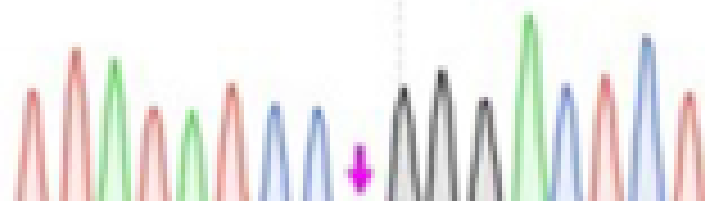

H 247

T T A T A T C C **R** G G G A C T C T
[A/G]

H 251

T T A T A T C C **R** G G G A C T C T
[A/G]

The tag SNP rs2455069-A>G in heterozygous mode of inheritance was revealed in all the affected individuals, and identified their haplotype by means of the 9 SNPs in the LD block: rs1697573, rs273634, rs1566576 rs12609179 rs273638, rs273639, rs1697553, rs3865444, and rs12459419 (Fig. 3).

FIGURE 3. *CD33* genomic structure. *CD33* genomic structure on chromosome 19q13.41. Most SNPs are located outside the coding sequence; the SNP rs3865444 is in the promoter region. Exon 2 harbors the tag SNP rs2455069.
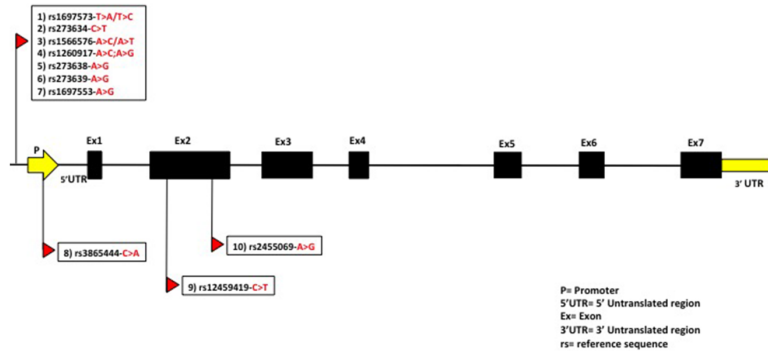

The genomic location where the block maps also comprises a region that includes the Siglecs, and is part of an immunoglobulin superfamily (IgSF) of cell surface and soluble proteins involved in the recognition, binding or adhesion processes of cells (Fig. 4). Based on the GeneHancer Regulatory Elements and Gene Interactions track http://genome.ucsc.edu/s/cocopow/hg38_GeneHancer, it appears that the *CD33* and *SIGLECL1* genes may have an interaction. Specifically, both GH19J051245 and GH19J051223 on chromosome 19q13.41could be promoters/enhancers that regulate and interact with the *CD33* and *SIGLECL1* genes (https://www.genecards.org/cgi-bin/carddisp.pl?gene=SIGLECL1&keywords=GH19J051223) https://www.genecards.org/cgi-bin/carddisp.pl?gene=SIGLECL1(Fig. 4).

FIGURE 4. Chromosome 19 q13.41 region. The upper part shows *CD33* and all 9 *SIGLEC* genes in the IgSF region located in 19q13.41. The bottom part shows the GeneHancer interaction between *CD33* and *SIGLECL1* genes and regulatory elements GH19J051223 and GH19J051245.
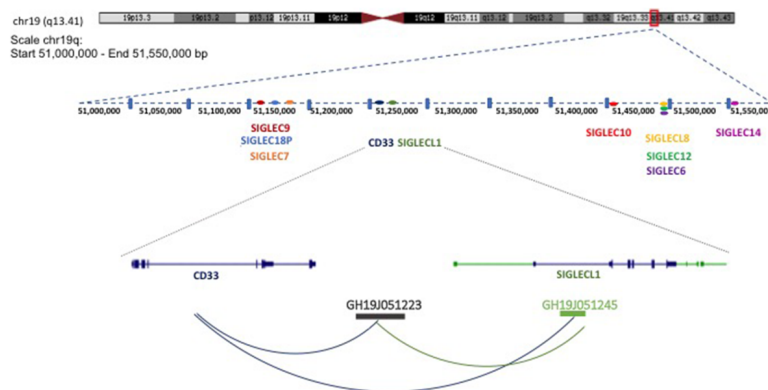
We tested the expression of both *CD33* and *SIGLECL1* in specific samples in cases where biological material was available. In particular, we analyzed the gene expression for D8-H118 and 2 random AD samples, H274 and H299, which had the heterozygote rs2455069 SNP allele (Table 3) and (Fig. 5A, B). The primers used for the analysis are specified in Supplementary Data Table S4.

FIGURE 5. *CD33* and *SIGLECL1* mRNA qPCR. qPCR of **(A)** *CD33* and **(B)** *SIGLECL1* genes (RefSeq NM_173635) on affected individual H188 who is part of the D8 disease-bearing family carrying the tag SNP. AD samples (H274; H299) carrying the tag SNP and healthy controls are represented as the mean of 3 individuals.

A CD33
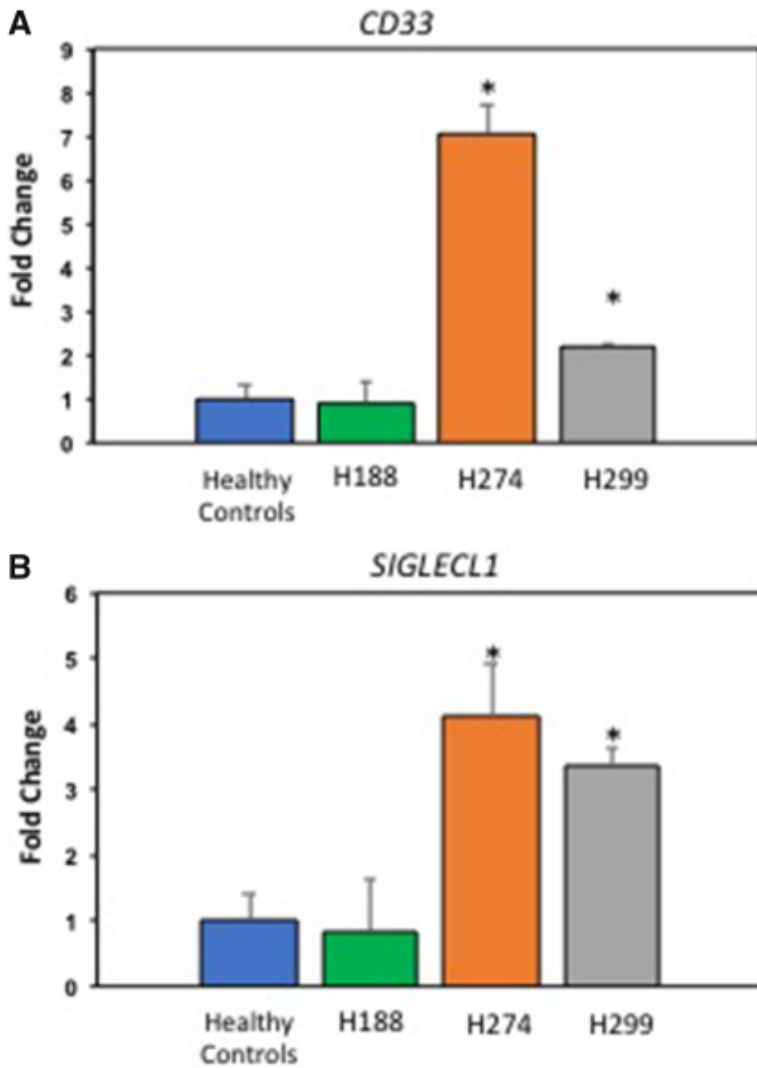
B SIGLECL1

Our results demonstrate increased expression of the *CD33* and *SIGLECL1* genes in both of the AD samples, H274 and H299, while the affected individual D8-H188 showed no variation in gene expression (Fig. 5A, B).

Alternative splicing mechanisms for the *CD33* gene generate 2 different isoforms, a full-length (*CD33M*) and a truncated isoform (*CD33m*) lacking exon 2. To assess the expression pattern of these 2 different *CD33* isoforms in patients belonging to the same family with the unusual phenotype, we performed qPCR on the RNAs extracted from leucocytes and compared the ratios between *CD33M* and *CD33m* isoforms (Fig. 6A–C). We also analyzed the total *CD33* expression in all of the family members, the healthy controls and the individuals affected by AD (Fig. 6B, C). When we compared the data from all of the healthy controls with those obtained from affected individuals, our results demonstrated an imbalance between the 2 isoforms. The ratio between normal and affected individuals was decreased because of the higher expression of *CD33m* in the AD individuals and in H217, the affected daughter, compared with the healthy controls. This observation likely represents an unusual occurrence. In contrast, we observed an increase in the ratios in H252, the healthy mother, in H251, the affected father, and H247 and H188, the affected sons (Fig. 6A). We note that we detected an ∼2.5-fold higher level of total *CD33* expression in H217 compared with healthy controls, while AD patients displayed a 6-fold increased expression of the total message (Fig. 6B).

FIGURE 6. **(A–C)** Alternative splicing and total expression of *CD33*. **(A)** mRNA ratio of alternative splicing expression of 2 *CD33* isoforms, a full-length (*CD33M*) and a truncated (*CD33m*) isoform lacking exon 2. **(B)** Total *CD33* expression showing an

increased level of ~2.5-fold in H217 compared with the healthy controls and a 6-fold increase in AD patients. **(C)** Table first column: ID patients and control; second column: Genotype Tag SNP rs2455069 (A>G; MAF: G = 0.335064/1678-1000Genomes). **(A)** and **(B)** show analyses in a pool of healthy controls, a pool of AD samples (H274; H299) carrying the tag SNP, and D8 family patients H247, H217, H188, and H251, and H252, the healthy mother.
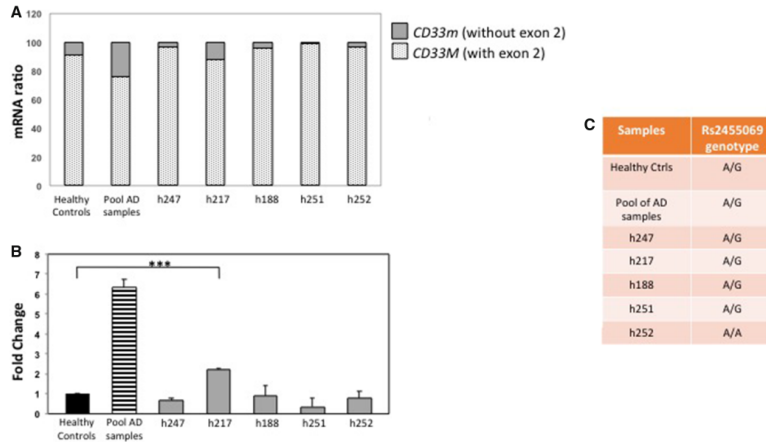
DISCUSSION

As a result of our clinical screening we identified an unusual family, D8 (Table 2; Fig. 1), carrying a behavioral phenotype with predominant executive/attention and memory problems that progressed rapidly and also evolved into episodic memory disturbances and dementia. According to this cognitive phenotype, FDG-PET at early stages showed a common pattern of relative hypometabolism involving the frontal cortex, typically reported in bvFTD, and in the superior parietal cortex/precuneus known to be involved early in AD (18). Interestingly, all patients displayed cognitive and/or cerebral cortical metabolic profiles in part shared bvFTD and AD. The lack of increased amyloid load, at least in 2 of our patients, is more suggestive of FTD rather than AD neuropathological substrate (19). We analyzed these patients for FTD and AD candidate genes, including *GRN*, *MAPT*, and *APOE*. For the father, we additionally performed an analysis *C9orf72*, which excluded the presence of triplet expansion. Our results identified the presence of a *CD33* SNPs block segregating in all the affected individuals of the family, previously described to be associated with cognitive decline (10). The exclusion of *GRN* mutation is of relevance since a cognitive/metabolic profile involving both anterior and posterior cortical regions has been described (3). The LD block was segregating with the affected individuals in the identified family (Table 2; Figs. 1 and 2) and is located in a genomic region upstream of the *CD33* gene. Although our analyses show an inconsistent direct expression effects of the LD block across the different patients, we cannot exclude the possibility that the unusual phenotype could be the result of other nonidentified isoforms that will surface when we analyze total gene expression. In addition, we cannot exclude the presence of epigenetic effects, such as different methylation patterns, which may further affect the expression profiles. Further investigation will be required to assess these hypotheses. Our analysis shows that the block is located inside a Siglec region, leading to the hypothesis of its involvement in the FTD and AD pathogenesis. This large IgSF may have a role in the development of the clinical appearance of cognitive decline in our patients (Figs. 4 and 5A, B).

Genome-wide association studies identified microglial receptor CD33 as being involved in dementia. Most likely, the *CD33* gene underwent comparable evolutionary changes and developed alleles involved with risk factors causing detrimental effects on cognition. Increased levels of proinflammatory mediators have been measured in the brains and cerebrospinal fluid of AD patients (20). This could be indicative of an AD network based on activated microglia, the primary cells of innate immunity in the central nervous system. IL-1β, IL-6, TGF-β, and several other cytokines have been found to accumulate around amyloid plaques (Aβ) in the brains of AD patients (21) and chronic microglial activation may contribute to the development and progression of neurodegenerative disorders. The roles of microglia also include the phagocytosis of cellular debris and the uptake and clearance of Aβ (22, 23). Activated microglial cells, which can have either neurotoxic or neuroprotective effects, express various membrane receptors. We have already demonstrated that the circulating levels of IL-1 family cytokines and receptors are significantly increased in AD patients (6). In particular, IL-1Rα, sIL-1R1, sIL-1R3, sIL-1R4, and IL-18BP have characteristic profiles in AD patients, providing support for the hypothesis that inflammation is a factor of AD (6, 24–26) and that memory features may be

regulated by it. This implies that active and persistent inflammation biomarkers exist at the tissue level, to which the organism may respond by activating multiple control mechanisms, albeit inefficiently.

Two *CD33* SNPs were described to be associated with AD (8) and an alternate splice pattern was hypothesized to affect the preservation of a cell's ability to take up and clear Aβ. The alternatively spliced *CD33* isoform lacks the IgV-set domain encoded by exon 2, which is implicated in CD33 sialic-acid binding capacity. CD33 in turn modulates the uptake of Aβ plaques, and the interaction of CD33 with sialic acids is crucial to permit this effect (7). The protective effect may lead to the preservation of the Aβ clearance functionality (27). Furthermore, it is clear that interactions of CD33 with sialic acids are crucial for this uptake, and that aberrant CD33 molecules have an effect on memory (24, 28). In microglia, CD33 binds extracellular sialylated glycans on other cells or pathogens and its cytoplasmic domain signals via phosphatidyl-inositol-3 kinase (PI3K) to dampen microglial phagocytosis (20). Our data confirm the expression of 2 *CD33* isoforms, which in combination, might be linked with the disease phenotype (Fig. 4A, B) (7, 29). The 2 splice variant transcripts could represent a full-length variant associated with the "AD-risk" (allele *CD33M*) and a shorter "AD-protective" isoform (allele *CD33m*). The mechanism by which the *CD33m* isoform increases β-amyloid clearance remains unknown. However, it is likely that the *CD33m* isoform protective properties are linked to a loss of function (29). It is possible that the lack of cerebral cortical amyloid load found in 2 of our patients might be related to the AD-protective role of *CD33m* isoform. This hypothesis should be further investigated in a larger series of patients using additional biomarkers.

The *CD33*-like locus on chromosome 19q13.4 includes 8 *Siglec* genes and 2 *Siglec* pseudogenes (Fig. 2) (30). The proteins encoded by all *Siglec* genes in this locus consist of a signal peptide and one V-set immunoglobulin domain, with the exception of the long form of the *SIGLECL1* gene, which has an extra V-set domain (31) followed by a variable number of C2-set immunoglobulin domains (32). Interestingly, we identified a tag SNP in *CD33* exon 2, rs2455069-A>G that leads to the change Arg69Gly in the CD33 amino acid sequence, which in turn leads to a change of allosteric forms that may compromise the interaction with sialic acids (data not shown) with a putative strong effect on protein function. This may compromise the interaction with sialic acids as part of the LD block associated with the disease in our family. In general, familial association results are easier to analyze given the reduced genomic heterogeneity compared with sporadic patients. The block crosses the IgSF cluster region and has a potential regulatory function on gene expression of different Siglec member isoforms. *CD33* and *SIGLECL1* exhibit a significantly increased expression level in AD and, to a lesser extent, in the patient family as a consequence of the young age and to a mixed expression of the memory phenotype. Analysis by GeneHancer (http://genome.ucsc.edu/s/cocopow/hg38_GeneHancer) showed a putative shared expression regulatory mechanism by at least 2 promoters/enhancers (Fig. 4). GH19J051223 represents a promoter for *SIGLECL1*, while GH19J051245 may have some interactions with the *SIGLECL1* gene and a role in regulating *CD33* expression, leading to the hypothesis of a common regulatory pathway.

In conclusion, we infer the existence of expression and a functional relationship between at least the 2 genes *CD33* and *SIGLECL* that may be relevant to the pathogenesis of FTD and AD.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. Niu H, Álvarez-Álvarez I, Guillén-Grima F, et al Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. Neurologia 2017;32:523–32

2. Onyike CU, Diehl-Schmid J. The epidemiology of frontotemporal dementia. Int Rev Psychiatry 2013;25:130–7

3. Milan G, Napoletano S, Pappatà S, et al GRN deletion in familial Frontotemporal dementia showing association with clinical variability in 3 familial cases. Neurobiol Aging 2017;53:193

4. Cairns NJ, Bigio EH, Mackenzie IRA, et al Neuropathologic diagnostic and nosologic criteria for frontotemporal lobar degeneration: Consensus of the Consortium for Frontotemporal lobar degeneration. Acta Neuropathol 2007;114:5–22

5. Warren JD, Rohrer JD, Rossor MN. Clinical review. Frontotemporal dementia. BMJ 2013;347:f4827

6. Italiani P, Puxeddu I, Napoletano S, et al Circulating levels of IL-1 family cytokines and receptors in Alzheimer's disease: New markers of disease progression? J Neuroinflammation 2018;15:342

7. Griciuc A, Serrano-Pozo A, Parrado AR, et al Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. Neuron 2013;78:631–43

8. Malik M, Simpson JF, Parikh I, et al CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. J Neurosci 2013;33:13320–5

9. Bhattacherjee A, Rodrigues E, Jung J, et al Repression of phagocytosis by human CD33 is not conserved with mouse CD33. Commun Biol 2020;3:36

10. Nettiksimmons J, Tranah G, Evans DS, et al Gene-based aggregate SNP associations between candidate AD genes and cognitive decline. Age (Dordr) 2016;38:41

11. Ferrari R, Mok K, Moreno JH, et al Screening for C9ORF72 repeat expansion in FTLD. Neurobiol Aging 2012;33:1850.e1–11

12. Renton AE, Majounie E, Waite A, et al A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011;72:257–68

13. Blauwendraat C, Faghri F, Pihlstrom L, et al NeuroChip, an updated version of the NeuroX genotyping platform to rapidly screen for variants associated with neurological diseases. Neurobiol Aging 2017;57:247

14. Graf P, Uttl B, Tuokko H. Color- and picture-word Stroop tests: Performance changes in old age. J Clin Exp Neuropsychol 1995;17:390–415

15. de Mendonça A, Ribeiro F, Guerreiro M, et al Frontotemporal mild cognitive impairment. J Alzheimers Dis 2004;6:1–9

16. Shin MS, Park SY, Park SR, et al Clinical and empirical applications of the Rey-Osterrieth complex figure test. Nat Protoc 2006;1:892–9

17. Aprahamian I, Martinelli JE, Liberalesso Neri A, et al The clock drawing test: A review of its accuracy in screening for dementia. Dement Neuropsychol 2009;3:74–81

18. Brown RK, Bohnen NI, Wong KK, et al Brain PET in suspected dementia: Patterns of altered FDG metabolism. Radiographics 2014;34:684–701

19. Jack CR Jr, Bennet DA, Blennow K, et al NIA-AA research framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement 2018;14:535–62.

20. Rosenthal SL, Kamboh MI. Late-onset Alzheimer's Disease genes and the potentially implicated pathways. Curr Genet Med Rep 2014;2:85–101

21. McGeer PL, McGeer EG. Inflammation and the degenerative diseases of aging. Ann N Y Acad Sci 2004;1035:104–16

22. Heneka MT, Carson MJ, El Khoury J, et al Neuroinflammation in Alzheimer's disease. Lancet Neurol 2015;14:388–405

23. Prinz M, Priller J, Sisodia SS, et al Heterogeneity of CNS myeloid cells and their roles in neurodegeneration. Nat Neurosci 2011;14:1227–35

24. Ridolfi E, Barone C, Scarpini E, et al The role of the innate immune system in Alzheimer's Disease and Frontotemporal lobar degeneration: An eye on microglia. Clin Dev Immunol 2013;2013:1–11

25. Brosseron F, Krauthausen M, Kummer M, et al Body fluid cytokine levels in mild cognitive impairment and Alzheimer's disease: A comparative overview. Mol Neurobiol 2014;50:534–44

26. Delaby C, Gabelle A, Blum D, et al Central nervous system and peripheral inflammatory processes in Alzheimer's disease: Biomarker profiling approach. Front Neurol 2015;6:181

27. Hesse R, Wahler A, Gummert P, et al Decreased IL-8 levels in CSF and serum of AD patients and negative correlation of MMSE and IL-1beta. BMC Neurol 2016;16:185

28. Raj T, Ryan KJ, Replogle JM, et al CD33: Increased inclusion of exon 2 implicates the Ig V-set domain in Alzheimer's disease susceptibility. Hum Mol Genet 2014;23:2729–36

29. Heneka MT, Kummer MP, Latz E. Innate immune activation in neurodegenerative disease. Nat Rev Immunol 2014;14:463–77

30. Siddiqui SS, Springer SA, Verhagen A, et al The Alzheimer's disease-protective CD33 splice variant mediates adaptive loss of function via diversion to an intracellular pool. J Biol Chem 2017;292:15312–20

31. Yousef GM, Ordon MH, Foussias G, et al Genomic organization of the siglec gene locus on chromosome 19q13.4 and cloning of two new siglec pseudogenes. Gene 2002;286:259–70

32. Foussias G, Taylor SM, Yousef GM, et al Cloning and molecular characterization of two splice variants of a new putative member of the Siglec-3-like subgroup of Siglecs. Biochem Biophys Res Commun 2001;284:887–99

AUTHOR QUERIES

Query: AQ1: Please check all author names and affiliations. Please check that author surnames have been identified by a pink background in the PDF version, and by green text in the html proofing tool version (if applicable). This is to ensure that forenames and surnames have been correctly tagged for online indexing.
Author Response: Accept

Query: AQ2: Please provide the division/department/university details for the affiliations where applicable.
Author Response: for Institute of Biochemistry and cell Biology(IBBC) add : National Research Council of Italy (CNR) Naples Italy;also University of Naples Federco II add: Naples, Italy; Institute of Genetics and Biophysics (IGB)add: National Research Council of Italy (CNR) Naples Italy;also for Geriatric Center "Frullone" ASL NA1, add : Naples Italy.department of internal medicine & surgery university of Naples Federico II add: Naples Italy

Query: AQ3: Please provide complete postal address with postal code for the Correspondence section.
Author Response: NeurOMICS Lab National Research Council of Italy (CNR)Institute of Biochemistry and Cell Biology (IBBC) via Pietro Castellino 111, 80131 NaplesItaly

Query: AQ4: Please confirm if the footnote 'Emilia Vitale is responsible…' is OK as placed.
Author Response: Accept

Query: AQ5: Please check that funding is recorded in a separate funding section if applicable. Use the full official names of any funding bodies, and include any grant numbers.
Author Response: for COINOR change in : COINOR grant STAR linea1-2018 (18-CSP-UNINA-042)for POR change to :P.O.R. Campania FESR 2007-2013 Project MOVIE (CUP B25C13000240007)

Query: AQ6: You may need to include a "conflict of interest" section. This would cover any situations that might raise any questions of bias in your work and in your article's conclusions, implications, or opinions. Please see https://academic.oup.com/journals/pages/authors/authors_faqs/conflicts_of_interest.
Author Response: Accept

Query: AQ7: Please check that all web addresses cited in the text, footnotes and reference list are up-to-date, and please provide a 'last accessed' date for each URL.
Author Response: Accept

Query: AQ8: As per journal style, genes and mutants should be in italics (and proteins should be in roman). Please check and make modifications as necessary.

Author Response: ABSTRACT PAGE : CD33 line 4 page should be italics.INTRODUCTION: line 29 other genes related to CD33..... Siglecs in italics.line 35 change: The mouse Cd33 and in italicshuman CD33 in italicsline 36 ...the importance of studying human ..change CD33 in italics

Query: AQ9: If your manuscript has figures or text from other sources, please ensure you have permission from the copyright holder. For any questions about permissions contact jnls.author.support@oup.com.
Author Response: no figures were taken from other sources

Query: AQ10: Please provide the significance for the bold entries in Table 1. Also, please confirm if the definitions of columns 1, 2, etc. are OK as set in the footnote of Tables 1–3.
Author Response: the bold was only for aesthetics, can be removed; you can add GENE Name in the definition for column 1; the column 2, 4 and 5 are not aligned words are overlapping; as well as lines PSEN2, TREM2, MAPT.

ATTACHMENTS

A1 File: 1_CD33 and SIGLEC Supplementary Data accepted 5-7-20 table 1[AU].docx has been attached by Author
A2 File: 1_CD33 and SIGLEC Supplementary Data accepted 5-7-20 table 1[AU].pdf has been attached by Author